DOCUMENT RESUME

ED 072 479 CS 500 126

AUTHOR Becker, Samuel L.
TITLE Evaluation of Academic Programs.
PUB DATE Dec 72
NOTE 12p.; Paper presented at the Annual Meeting of the Speech Communication Assn. (58th Chicago, December 27-30, 1972)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Academic Education; Academic Performance; Academic Standards; *College Faculty; *Communication (Thought Transfer); Cost Effectiveness; *Departments; *Educational Accountability; *Evaluation; Evaluation Methods; Organization

ABSTRACT

Recent developments in faculty organization and accountability are discussed separately in terms of their implications for evaluation of academic programs and departments. Faculty organization is examined in relation to each components' (individual faculty members, departments, colleges, and other assorted units) accessibility to and utilization of feedback data assessing their performance and effectiveness. Recommendations for checking the possible inhibiting effects of departmental autonomy upon department program evaluation are advanced. Suggestions for a nation-wide effort (coordinated by national organizations) to establish goal consensus and accountability criteria among institutions according to field, are made. Through organizations such as the Speech Communication Association and Association of Departments and Administrators in Speech Communication, the general public and decision-makers in institutions of higher learning can be educated about the value of increasing the communication skills and sophistication of individuals, groups, and organizations. (Author/LG)

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 072479

EVALUATION OF ACADEMIC PROGRAMS

Samuel L. Becker
University of Iowa

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Samuel L. Becker

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

Every parent who has ever managed to live through the teen-age years of ~~their~~ his children is more than familiar with the evanescent norms that shape their children one way one week, and another way the next. We joke and laugh about these teen-age fads--whenever we can do so through our tears, that is.

Such fads that sweep the teen-age community are cause enough to weep. Even greater cause for many of us, though, are the fads that almost as regularly sweep the educational community. Just in recent times, among others, we have had educational television, then programed instruction, then behavioral objectives, and now "accountability" or, as it more often manifests itself in the public schools, "performance contracting."

Though I have observed the course of teen-age fads over many years, I have not yet discovered the motive forces--those forces that give rise to them, that mark their peak, and that signal their decline. With educational fads, on the other hand, I think that I can observe some patterns. A strong motive force behind the rise of many of them is the push for efficiency, for educating more students or educating them better with less money. As each fad approaches its peak, we can usually see a plethora of studies purporting to demonstrate the usefulness of the particular practice. This is followed by the realization that most of the studies have not been well done and that, therefore, their results are questionable. So well-done, large-scale studies follow, and the results

AS 500 126

of these usually mark the beginning of the decline. Such a large-scale study may well have just been done on performance contracting. After a great many small demonstration studies, the United States Office of Economic Opportunity was so confident of the success of performance contracting that it spent over six-million dollars in a large, nation-wide study of performance contracting with under-achieving students. Thirty-one companies bid for the opportunity to participate, even though they would be paid only for each student who was able to gain one school year of growth in reading and arithmetic during the ~~one~~ 2 school year. Half or more of those companies that participated are now out of business. The pupils in the experiment fell far short of the goal. Even more important, when the gains on these youngsters were compared to the gains of an equal number of comparable pupils in control groups, no effects of performance contracting could be detected.[1]

I am not suggesting here that all of these fads are worthless or that we ignore them. Obviously, many have within them the seeds for new blossoms in the educational garden--new growth that can contribute to the achievement of our varied missions. What I am suggesting is that we must avoid both kinds of signal responses which some members of the academic community consistently make to new developments. Some of us tend habitually to pooh-pooh anything new--asserting either that it is not, in fact, new at all or else that it is useless or even destructive. Others of us tend, just as habitually, to immediately gather every new development to our institutional bosoms, either because we perceive that government and foundation grants tend to go to innovators or else because we are habitually dissatisfied with what we are doing and so are always ready to try something

new. I hope that we here can find some middle ground--that we can analyze these new developments to discover the potential strengths and weaknesses and to develop means of capitalizing on the former while minimizing the latter. This is the framework within which I will be discussing faculty organization, accountability, and the evaluation of academic programs.

The recent developments in faculty organization and in accountability which we have been discussing at this conference have quite different sorts of implications for the evaluation of academic programs or of departments. Therefore, I will consider them separately; first faculty organization and then accountability.

FACULTY ORGANIZATION

We have talked at this conference primarily about one type of faculty organization, the union or something which approaches the union. I believe that we should also consider another type of faculty organization, that within each institution-- the organization into departments, colleges, and other assorted units. Both of these kinds of organization can affect the degrees of freedom that we have for an adequate evaluation of academic programs and departments.

As teaching unions gather strength, they may inhibit or attempt to inhibit the gathering of some of the kinds of data that are needed for an adequate assessment of a department's affectiveness. Since that effectiveness is, in large part, the cumulative effectiveness of the faculty members within the department, we must have the freedom to assess as completely as we can the effectiveness of each individual faculty member. We must be able to get

student ratings of each individual teacher, as well as student ratings of courses and programs. We must also attempt to get some comparative data on the performance of students who have worked under each faculty member. (I will return to this point later when I discuss accountability.) Any restrictions which union contracts place upon the gathering and use of such data will limit the validity of departmental evaluations.

Departmental and college structures may inhibit such evaluations in another way. Though I am a fervent believer in departmental autonomy, as I assume you are, we must recognize that such autonomy may eliminate what ought to be some of the major criteria in the evaluation of a department. One such criterion is the degree to which a department's course offerings serve the students of other departments and colleges, as opposed to serving only or primarily its majors. The other related criterion is how well a department uses what is available from other departments. We must get away from the sort of departmental autonomy which leads us to believe that only the faculty in a department ~~is~~ are qualified to determine what a department ought to teach or who we ought to teach it to. Our departments of speech communication and theatre ought to be evaluated not only on the basis of the evidence of how effectively we are meeting the learning needs of our majors, but also whether we are meeting the learning needs of students in American Civilization (with our Public Address courses), of students in pre-law(with our Argumentation and Persuasion courses), of students in English (with our dramatic literature and Rhetorical Theory courses), and students in Journalism and Art (with our broadcasting and film courses), etc. We ought also to be evaluated

on how well we use the courses from other departments to educate our students more effectively and efficiently. We should be asked whether we insist on teaching our own statistics courses when there are better statistics courses taught by the department of Psychology or Mathematics, or our own historiography courses when they can be better taught by a specialist in the History department, etc. When we insist on teaching such courses when there are substantial courses in the departments which profess primary expertise in those areas, we are wasting our resources and not properly serving our students.

There is another aspect of departmental service which needs to be taken account of in any departmental evaluation. Not only must we assess the degree to which each department's offerings serve the needs of other departments and colleges; we must also consider the degree to which its offerings are appropriate for the student body and the region which it serves.[2] No longer can a department set its goals and determine its programs independent of community needs. If a region needs help with dialects or the development of skills in organizing and maintaining groups for community action programs, and we offer only oral interpretation, formal public speaking, and the history of rhetorical theory, serious questions should be raised about whether we deserve the financial support of that community.

ACCOUNTABILITY

Let us consider now some of the questions raised by the current push for accountability. The principle of accountability is good, one with which none of us can argue. Clearly, we in the

colleges and universities must be accountable to the public which supports us and to those agencies which stand between us and the public. The principle of accountability is fine. I am less sure about the practice. In principle, an accountability system measures the input to an institution (the dollars allocated) against the output (the amount of education). In practice, our measure of dollars seems to be pretty good, but I have serious reservations about the measure of education that we have been using. That measure, as far as I can see, appears to be student credit hours generated. The amount or quality of learning which each of those credit hours represents seldom if ever gets taken into account. Until these factors are taken into account, this method of assessing the work of colleges and universities or of departments is useless, if not seriously deleterious. A department or an institution is not like a department store where one's primary concern is rapid turnover and, hence, maximum profit (or minimum loss). Basing evaluations of departments or institutions solely, or even primarily, on the average cost of student credit hours generated encourages a variety of practices which, in the long run, will lower the quality of education. It will encourage a department to keep promotions and tenure to a minimum and have rapid turnover of faculty instead because senior faculty cost more money and don't generate any more credit hours than inexperienced instructors or assistant professors. It will discourage a department from offering independent study and small group seminars, or keeping ample office hours and carrying on research.

Accountability procedures will not encourage these counterproductive practices if we build in proper measures of educational

effectiveness--if we can successfully combine valid measures of effectiveness with our measures of efficiency. The problem is simply that up until now, no institution that I can discover has been successful in combining them.

The reason for this failure I believe is clear. We have been trying to measure the effectiveness of education at too macroscopic a level; we have been trying to find a single measure which can be applied to any of the graduates of an institution. Up until now--and I see no possibility for predicting anything different in the future--colleges and universities have been spectacularly unsuccessful in defining their institutional goals in a meaningful way--in a way that makes possible the assessment of the degree to which those goals have been met. Being realistic, such attempts can probably never succeed except in a highly specialized professional institution such as a law or medical school where at least some of the goals are clearly defined by the licensing examinations. Otherwise, especially in the liberal arts college, the educational goals of the various students and the various departments are so heterogeneous that the institutional goals have little value except to impress parents and potential donors.

Even when such substantive tests as the Area Tests of the Graduate Record Examination are used, the effects of a college are virtually impossible to determine. Part of the difficulty is due to differences in initial abilities of different college populations. However, even when something like the aptitude test of the College Entrance Examination Board or of the American College Testing program is used to control for initial differences, we can still detect little effect of different colleges. The correlations

found between these entrance examination tests and the area tests is very high--often in the .90s. This means that there is not much variability left that a college can influence.[3]

Because of these various problems of evaluating effectiveness at the moment, some observers have suggested that while we are working on this problem we go ahead and evaluate solely on the basis of economic efficiency. Off hand this seems reasonable, but there are underlying dangers. In striving for economic efficiency, we could destroy some of the major values of higher education long before we come up with satisfactory means of assessing those values so that we can work them into the formulae.

In spite of these arguments, which neither the general public nor the governing bodies of our institutions probably find as compelling as we do, the trend toward using accountability procedures for assessing institutions and parts of institutions will undoubtedly continue, if not accelerate. If we are to avoid or minimize the dangers in the system, if we do not want to live with cost per student credit hour as the major criterion of the quality of job we are doing, we must develop valid measures of learning which can be built in as major factors in the accountability formulae. Not only would such measures overcome the major criticisms of accountability procedures, they would be a positive gain, for they would give students, faculty, administrators, and the public a reasonable basis for assessing a department or institution.

In my opinion, we will not be able to develop valid measures of institutional effect if we continue to concentrate on the very general measures of the impact of colleges which have been attempted

in the past. It is probably even useless--at least at this point in time--to attempt to develop criteria for departments or programs. The futility of attempting this in our field, for example, is demonstrated by the failure of experts in the Educational Testing Service, even with the cooperation of subject-matter experts from our field, to develop a satisfactory area test for speech communication as part of the Graduate Record Examination. Therefore, instead of trying and failing at that task again, we ought to begin with the courses or clusters of courses which tend to be taught in most departments or institutions and develop tests for them which can be used in assessing the quality of teaching in different institutions.

The measures that I am advocating cannot be developed by college or university administrators, or by the accountants or Ph.D.'s in educational administration who are probably planning most of the accountability systems, even if they should be interested in doing so. Only those of us who profess expertise in the various subject matter areas can do this. And, if we are to do it meaningfully, it must be done on a national basis. This means that we here and our colleagues must do the job for the field of speech communication, and that we can only do the job well if we coordinate our efforts through an organization such as the Speech Communication Association or the Association for Departments and Administrators of Speech Communication.

Standardized tests of this sort should be developed for all of the key courses or clusters of courses where there is consensus among institutions on course goals. Where such agreement does not exist, we must work toward bringing it about.

Lest I seem overly optimistic, I would point out that I recognize the difficulties of getting agreement on course goals from different institutions--even in a field such as ours where we profess expertise in problem-solving discussion. Certainly our past efforts at assessing the effects of speech communication courses have not been very successful. Consider, for example, the research that we have done on courses in public speaking. We have studied their effect on personality, on critical thinking, on attitudes toward speaking, on anxiety while speaking, and--once in a while--even on skill at speaking. The picture that emerges from research on even this simple problem-assessing the success of a single performance course--is ambiguous.[4]

To develop decent tests, not only must we agree among ourselves on what the general goals of each course should be, but we must also translate our general goals into behavioral objectives. And here we encounter another problem. The major value of a good set of behavioral objectives is their high degree of specificity which makes them measurable. But this very value creates a danger. When we insist upon highly specific behavioral objectives, we are likely forcing ourselves to omit some important educational outcomes. For example, one author, sophisticated in educational evaluation, has even made this comment about the use of behavioral objectives in the mathematics program of the primary schools.

> The unfortunate consequence of this atomization is that the interrelatedness of math concepts is lost and the statement is a tedious list of very trivial low-level skills.[5]

If this is a problem with the behavioral goals for primary school mathematics, where there seem to be agreed upon behaviors which

pupils should learn, how much more of a problem it is bound to be at the college level, especially in a field such as ours.

In spite of these problems, with the push from an organization such as ours, I am confident that we can develop useful measures which, for comparative purposes, have a high degree of validity. Another of the functions of SCA or ADASC can be to help us gather data from a good sample of departments with similar missions in similar institutions. Using the data from each cluster of similar institutions, we can develop norms against which the effectiveness of the course or cluster in each department can be measured.

With these measures of individual courses and clusters of courses, we will ultimately be able to get back to the evaluation of total departments, and even of colleges and universities. At that point, though, we will be able to evaluate these larger structures not by using a common measure for all students within them, but rather by combining the results of those individual measures which are most appropriate for each student in each institution.[6]

After we have developed reliable and valid measures of effectiveness of courses, clusters of courses, or even departments, colleges, and universities, we will be almost home--though not quite in the door. We will still have the problem of bringing together our measures of effectiveness with the measures of efficiency which are now being refined in the accountability studies. Ultimately, we will come to the point where we must place a dollar value on various kinds and amounts of educational gain. That is, ultimately we must ask of any particular kind or any particular

amount of gain, "Is it worth the cost?" For example, how much money is it worth to increase the knowledge of 100 students about communication processes by 10 per cent? Or how much is it worth to decrease the speech anxiety of 20 students by 5 per cent? When we reach these kinds of questions, research and researchers can offer little help. These questions of relative value are ones that must be answered by the society. Hopefully, though, it will be an enlightened society. Here also there is a function for organizations such as the SCA and ADASC. These organizations must be concerned with helping their members to educate the general public and the decision-makers in institutions of higher learning about the value of increasing the communication skills and sophistication of individuals, groups, and organizations.

When we have achieved these ends that I have suggested, evaluation of our programs and our departments should be a great boon to us and to our field.

FOOTNOTES

1. Ellis B. Page, "How We All Failed in Performance Contracting," Educational Psychologist, IX (1972), 40-41.

2. Point suggested by Dewey B. Stuit, "Measuring the Quality of a College or University," in Proceedings of the 1961 Invitational Conference on Testing Problems, chairman Paul L. Dressel (Princeton: Educational Testing Service, 1961), p. 78.

3. Rodney T. Hartnett, Accountability in Higher Education (Princeton: College Entrance Examination Board, 1971), p. 16.

4. For details, see Samuel L. Becker, "Research on Speech Pedagogy," in Dimensions of Rhetorical Scholarship, ed. Roger E. Nebergall (Norman: University of Oklahoma, 1963), pp. 31-48.

5. Hartnett, p. 13.

6. The problem of standardizing these varied measures so that they are comparable will be relatively simple since we will be con-concerned with deviations from norms rather than absolute scores.